



wwPDB EM Validation Summary Report ⓘ

Jul 15, 2025 – 01:19 PM JST

PDB ID : 9KHY / pdb_00009khy
EMDB ID : EMD-62359
Title : Terminator and trunk structure of Escherichia phage Mu
Authors : Zhou, J.Q.; Liu, H.R.
Deposited on : 2024-11-11
Resolution : 3.40 Å(reported)

This is a wwPDB EM Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

EMDB validation analysis : **FAILED**
MolProbity : 4-5-2 with Phenix2.0rc1
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.44

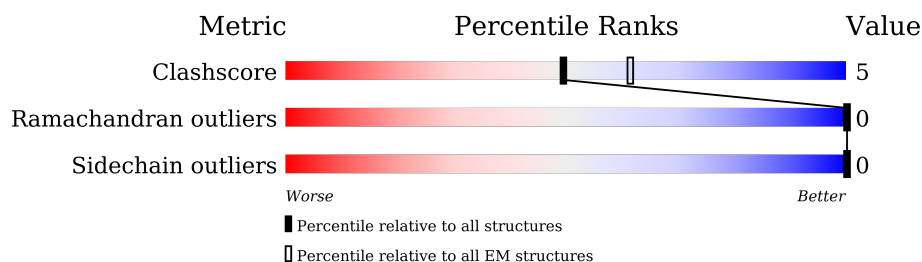
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.40 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



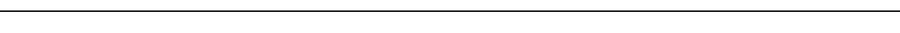
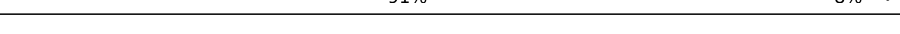
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$

Mol	Chain	Length	Quality of chain	
1	a	495	84%	15%
1	b	495	83%	16%
1	c	495	83%	17%
1	d	495	85%	15%
1	e	495	84%	16%
1	f	495	84%	16%
1	g	495	88%	12%
1	h	495	87%	13%
1	i	495	87%	12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	j	495	 87% 13%
1	k	495	 87% 13%
1	l	495	 88% 12%
2	1	182	 78% 18% .
2	2	182	 78% 18% .
2	3	182	 78% 18% .
2	4	182	 78% 18% .
2	5	182	 79% 17% .
2	6	182	 77% 18% .
3	A	118	 92% 7% .
3	B	118	 92% 7% .
3	C	118	 92% 8% .
3	D	118	 92% 8% .
3	E	118	 91% 8% .
3	F	118	 91% 8% .
3	G	118	 89% 10% .
3	H	118	 87% 12% .
3	I	118	 91% 8% .
3	J	118	 88% 11% .
3	K	118	 90% 9% .
3	L	118	 89% 10% .

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 63684 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Tail sheath protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	a	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	b	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	c	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	d	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	e	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	f	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	g	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	h	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	i	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	j	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	k	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		
1	l	494	Total	C	N	O	S	0	0
			3725	2343	639	729	14		

- Molecule 2 is a protein called Probable tail terminator protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	1	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		
2	2	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		
2	3	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
2	4	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		
2	5	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		
2	6	174	Total	C	N	O	S	0	0
			1384	879	243	257	5		

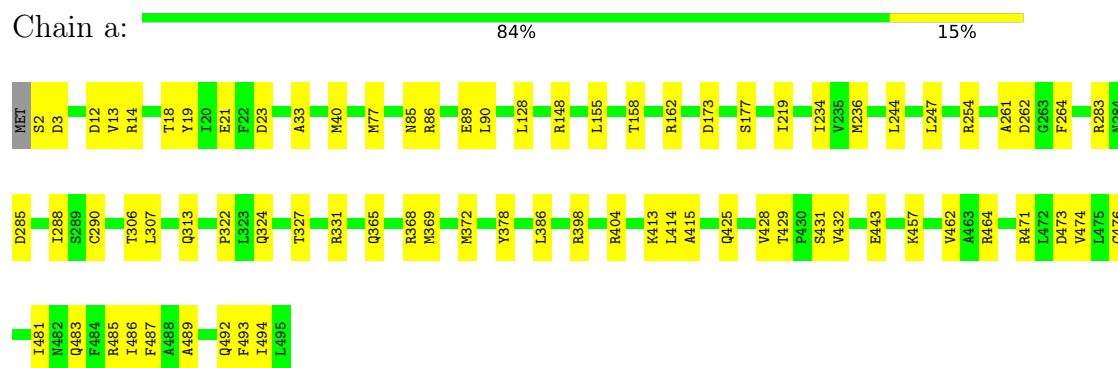
- Molecule 3 is a protein called Tail tube protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	G	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	H	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	I	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	J	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	K	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	L	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	A	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	B	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	C	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	D	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	E	117	Total	C	N	O	S	0	0
			890	555	156	174	5		
3	F	117	Total	C	N	O	S	0	0
			890	555	156	174	5		

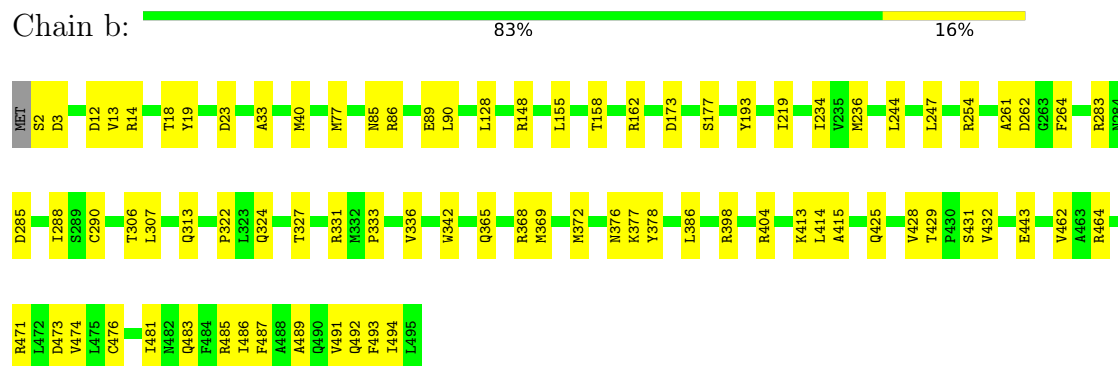
3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

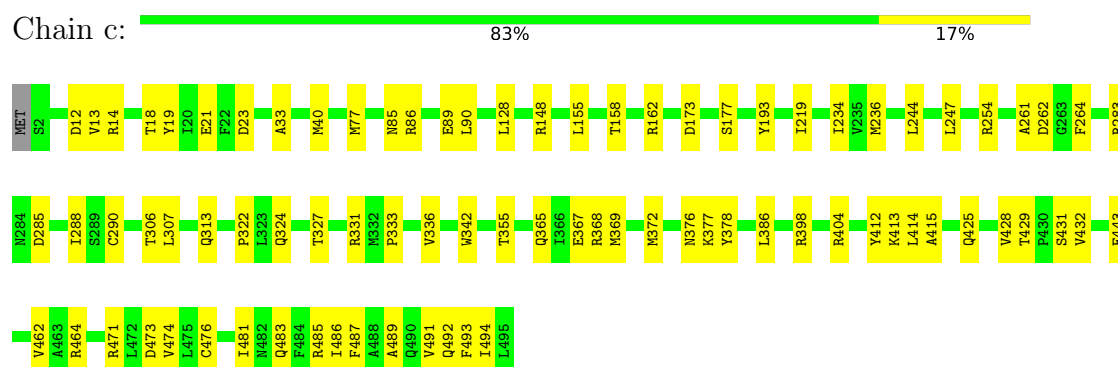
• Molecule 1: Tail sheath protein



• Molecule 1: Tail sheath protein

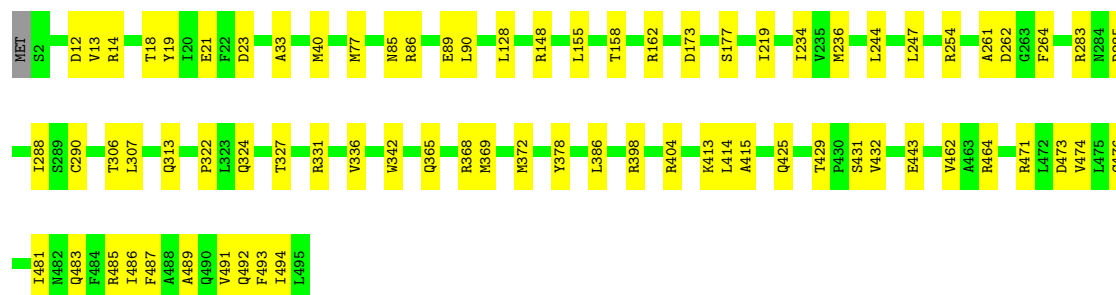


• Molecule 1: Tail sheath protein



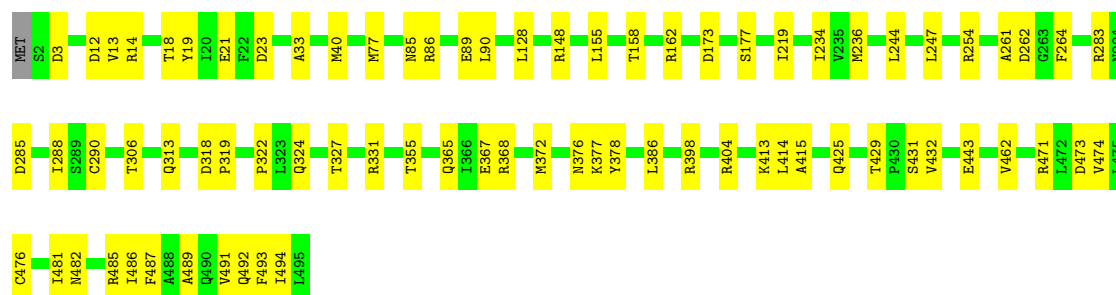
- Molecule 1: Tail sheath protein

Chain d:  85% 15%



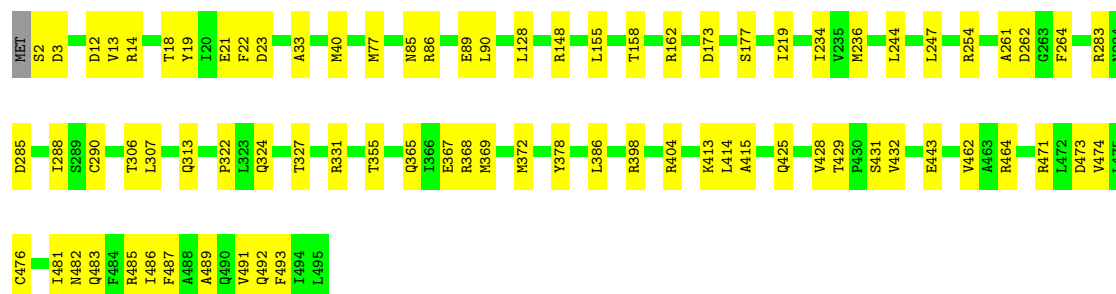
- Molecule 1: Tail sheath protein

Chain e:  84% 16%



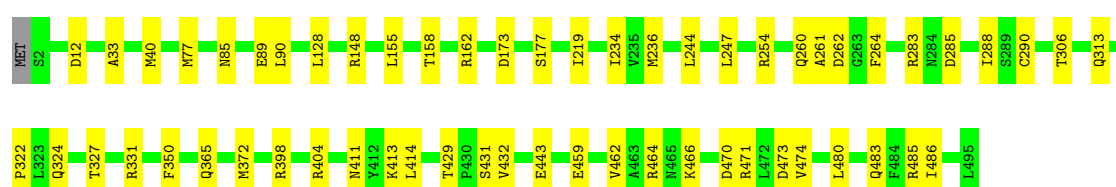
- Molecule 1: Tail sheath protein

Chain f:  84% 16%



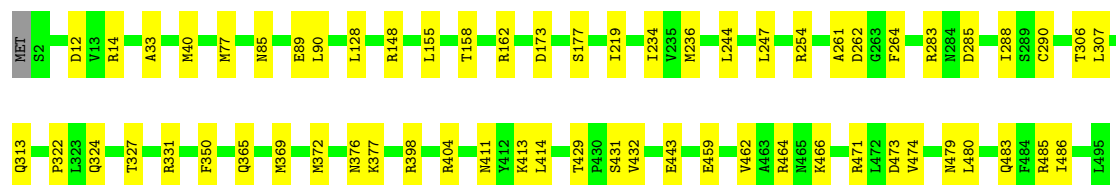
- Molecule 1: Tail sheath protein

Chain g:  88% 12%



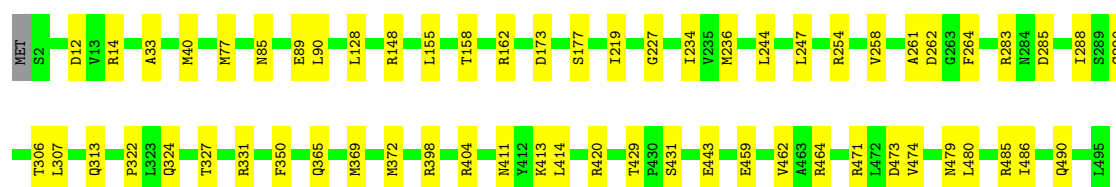
- Molecule 1: Tail sheath protein

Chain h:  87% 13%



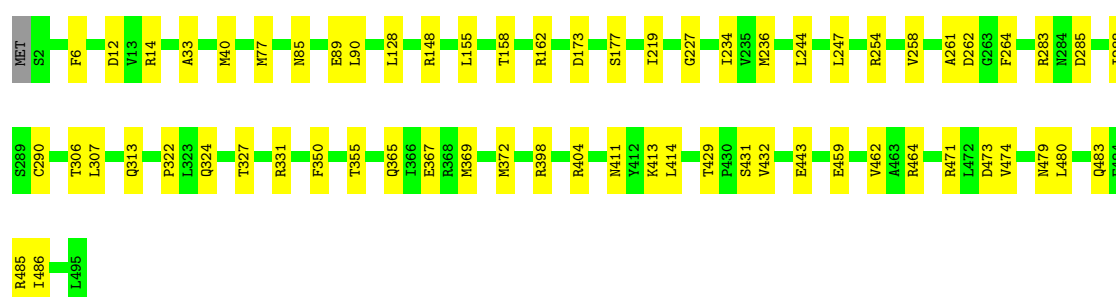
- Molecule 1: Tail sheath protein

Chain i:  87% 12%



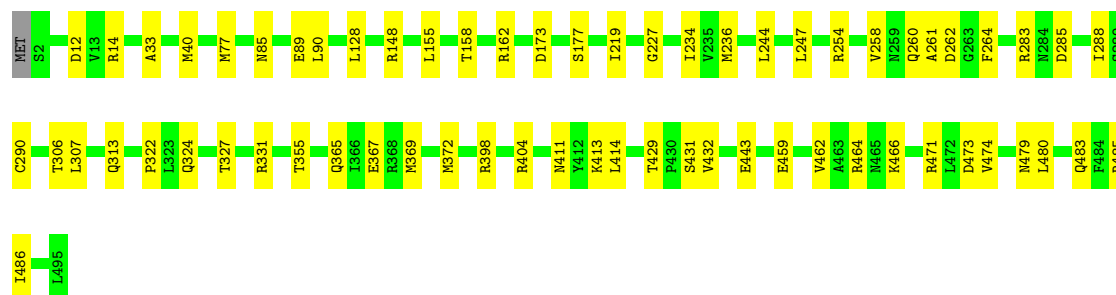
- Molecule 1: Tail sheath protein

Chain j:  87% 13%



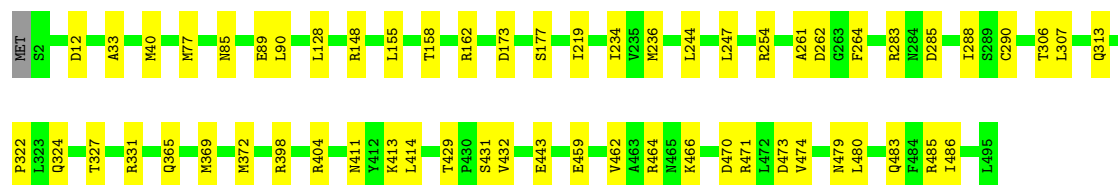
- Molecule 1: Tail sheath protein

Chain k:  87% 13%

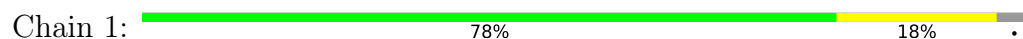


- Molecule 1: Tail sheath protein

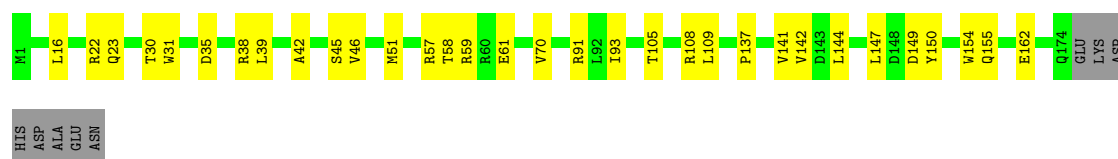
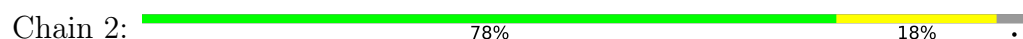
Chain l:  88% 12%



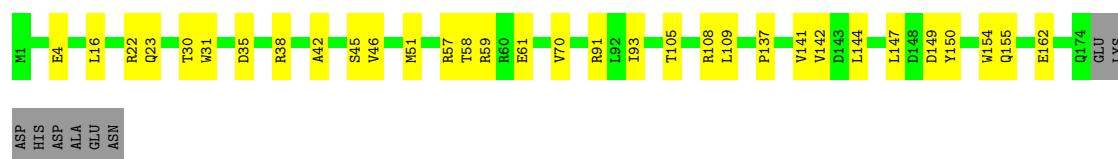
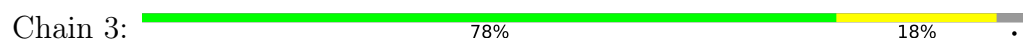
- Molecule 2: Probable tail terminator protein



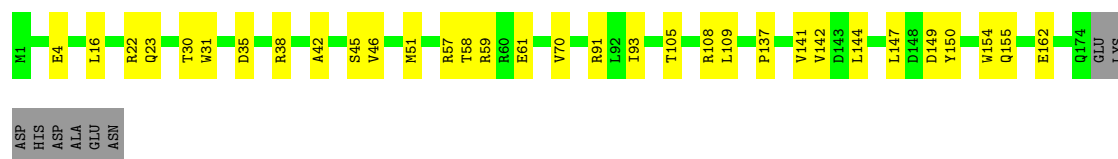
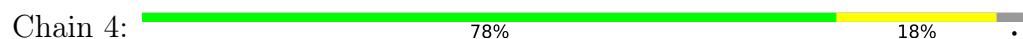
- Molecule 2: Probable tail terminator protein



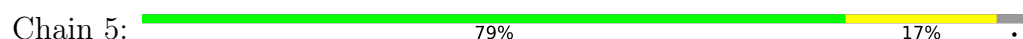
- Molecule 2: Probable tail terminator protein

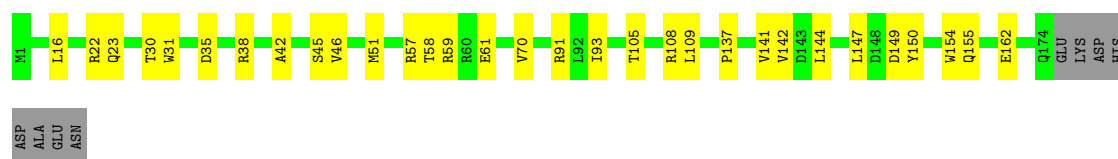


- Molecule 2: Probable tail terminator protein



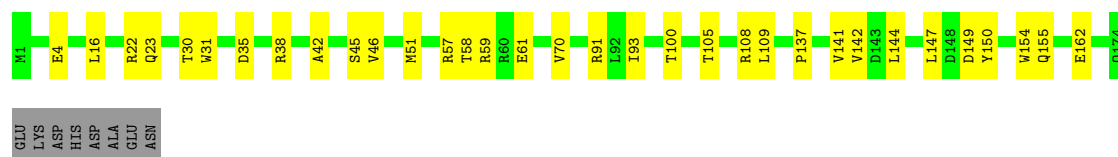
- Molecule 2: Probable tail terminator protein





- Molecule 2: Probable tail terminator protein

Chain 6: .



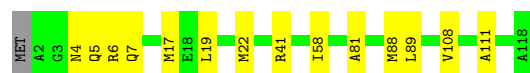
- Molecule 3: Tail tube protein

Chain G: .



- Molecule 3: Tail tube protein

Chain H: .



- Molecule 3: Tail tube protein

Chain I: .



- Molecule 3: Tail tube protein

Chain J: .



- Molecule 3: Tail tube protein

Chain K: .



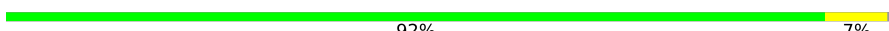
• Molecule 3: Tail tube protein

Chain L:  89% 10%

• Molecule 3: Tail tube protein

Chain A:  92% 7%

• Molecule 3: Tail tube protein

Chain B:  92% 7%

• Molecule 3: Tail tube protein

Chain C:  92% 8%

• Molecule 3: Tail tube protein

Chain D:  92% 8%

• Molecule 3: Tail tube protein

Chain E:  91% 8%

• Molecule 3: Tail tube protein

Chain F:  91% 8%

4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	55035	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	NONE	Depositor
Microscope	TFS KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	35	Depositor
Minimum defocus (nm)	2000	Depositor
Maximum defocus (nm)	4000	Depositor
Magnification	Not provided	
Image detector	GATAN K3 (6k x 4k)	Depositor

5 Model quality [i](#)

5.1 Standard geometry [i](#)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	a	0.10	0/3799	0.24	0/5174
1	b	0.10	0/3799	0.24	0/5174
1	c	0.10	0/3799	0.24	0/5174
1	d	0.10	0/3799	0.24	0/5174
1	e	0.10	0/3799	0.24	0/5174
1	f	0.10	0/3799	0.24	0/5174
1	g	0.10	0/3799	0.24	0/5174
1	h	0.10	0/3799	0.24	0/5174
1	i	0.10	0/3799	0.24	0/5174
1	j	0.10	0/3799	0.24	0/5174
1	k	0.10	0/3799	0.24	0/5174
1	l	0.10	0/3799	0.24	0/5174
2	1	0.11	0/1418	0.27	0/1935
2	2	0.11	0/1418	0.27	0/1935
2	3	0.11	0/1418	0.27	0/1935
2	4	0.11	0/1418	0.27	0/1935
2	5	0.11	0/1418	0.27	0/1935
2	6	0.11	0/1418	0.27	0/1935
3	A	0.11	0/906	0.27	0/1222
3	B	0.11	0/906	0.27	0/1222
3	C	0.11	0/906	0.26	0/1222
3	D	0.11	0/906	0.27	0/1222
3	E	0.11	0/906	0.27	0/1222
3	F	0.11	0/906	0.27	0/1222
3	G	0.11	0/906	0.26	0/1222
3	H	0.11	0/906	0.27	0/1222
3	I	0.11	0/906	0.27	0/1222
3	J	0.11	0/906	0.27	0/1222
3	K	0.11	0/906	0.27	0/1222
3	L	0.11	0/906	0.27	0/1222
All	All	0.10	0/64968	0.25	0/88362

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	a	3725	0	3651	64	0
1	b	3725	0	3651	68	0
1	c	3725	0	3651	66	0
1	d	3725	0	3651	61	0
1	e	3725	0	3651	63	0
1	f	3725	0	3651	64	0
1	g	3725	0	3651	39	0
1	h	3725	0	3651	43	0
1	i	3725	0	3651	41	0
1	j	3725	0	3651	43	0
1	k	3725	0	3651	41	0
1	l	3725	0	3651	40	0
2	1	1384	0	1349	25	0
2	2	1384	0	1349	25	0
2	3	1384	0	1349	27	0
2	4	1384	0	1349	25	0
2	5	1384	0	1349	27	0
2	6	1384	0	1349	28	0
3	A	890	0	871	5	0
3	B	890	0	871	5	0
3	C	890	0	871	7	0
3	D	890	0	871	6	0
3	E	890	0	871	7	0
3	F	890	0	871	7	0
3	G	890	0	871	9	0
3	H	890	0	871	11	0
3	I	890	0	871	8	0
3	J	890	0	871	10	0
3	K	890	0	871	10	0
3	L	890	0	871	9	0
All	All	63684	0	62358	692	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 692 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:d:322:PRO:HD2	1:k:411:ASN:HB2	1.62	0.81
1:c:322:PRO:HD2	1:j:411:ASN:HB2	1.66	0.78
1:a:476:CYS:SG	2:6:155:GLN:NE2	2.56	0.78
1:f:322:PRO:HD2	1:g:411:ASN:HB2	1.66	0.78
1:e:322:PRO:HD2	1:l:411:ASN:HB2	1.66	0.78

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	a	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	b	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	c	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	d	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	e	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	f	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	g	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	h	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	i	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	j	492/495 (99%)	480 (98%)	12 (2%)	0	100	100
1	k	492/495 (99%)	479 (97%)	13 (3%)	0	100	100
1	l	492/495 (99%)	480 (98%)	12 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	1	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
2	2	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
2	3	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
2	4	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
2	5	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
2	6	172/182 (94%)	167 (97%)	5 (3%)	0	100	100
3	A	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	B	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	C	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	D	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	E	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	F	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	G	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	H	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	I	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	J	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	K	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
3	L	115/118 (98%)	111 (96%)	4 (4%)	0	100	100
All	All	8316/8448 (98%)	8093 (97%)	223 (3%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	a	392/394 (100%)	392 (100%)	0	100	100
1	b	392/394 (100%)	392 (100%)	0	100	100
1	c	392/394 (100%)	392 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	d	392/394 (100%)	392 (100%)	0	100	100
1	e	392/394 (100%)	392 (100%)	0	100	100
1	f	392/394 (100%)	392 (100%)	0	100	100
1	g	392/394 (100%)	392 (100%)	0	100	100
1	h	392/394 (100%)	392 (100%)	0	100	100
1	i	392/394 (100%)	392 (100%)	0	100	100
1	j	392/394 (100%)	392 (100%)	0	100	100
1	k	392/394 (100%)	392 (100%)	0	100	100
1	l	392/394 (100%)	392 (100%)	0	100	100
2	1	148/155 (96%)	148 (100%)	0	100	100
2	2	148/155 (96%)	148 (100%)	0	100	100
2	3	148/155 (96%)	148 (100%)	0	100	100
2	4	148/155 (96%)	148 (100%)	0	100	100
2	5	148/155 (96%)	148 (100%)	0	100	100
2	6	148/155 (96%)	148 (100%)	0	100	100
3	A	91/92 (99%)	91 (100%)	0	100	100
3	B	91/92 (99%)	91 (100%)	0	100	100
3	C	91/92 (99%)	91 (100%)	0	100	100
3	D	91/92 (99%)	91 (100%)	0	100	100
3	E	91/92 (99%)	91 (100%)	0	100	100
3	F	91/92 (99%)	91 (100%)	0	100	100
3	G	91/92 (99%)	91 (100%)	0	100	100
3	H	91/92 (99%)	91 (100%)	0	100	100
3	I	91/92 (99%)	91 (100%)	0	100	100
3	J	91/92 (99%)	91 (100%)	0	100	100
3	K	91/92 (99%)	91 (100%)	0	100	100
3	L	91/92 (99%)	91 (100%)	0	100	100
All	All	6684/6762 (99%)	6684 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	h	192	ASN
1	l	43	GLN
1	i	43	GLN
1	j	192	ASN
3	A	5	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.